Architecting Storage Applications for the Public Cloud Economy

SDC EMEA 2020 - Tel Aviv

Josh Salomon
Senior Principal Software Engineer, Red Hat
jsalomon@redhat.com
Storage Applications for the Public Cloud

Agenda:
1) Why do we need Storage Applications in the public cloud
2) The public cloud structure and economy
3) Tips for developing economic Storage Applications for the public cloud.
Once Upon a Time in Windows 3.x Era

What is the difference?

while (GetProcAddress (&uMsg, NULL, 0, 0) > 0) {
    TranslateMessage (&uMsg);
    DispatchMessage (&uMsg);
}

while (WM_QUIT != uMsg.message) {
    if (GetProcAddress (&uMsg, NULL, 0, 0, PM_REMOVE) > 0) {
        TranslateMessage (&uMsg);
        DispatchMessage (&uMsg);
    } else if (IsBGWork()) {
        DoBGWork();
    }
}

*GetMessage is battery friendly while PeekMessage is not*
Storage Applications for the Public Cloud

- Why do we need Storage Applications in the public cloud
  - Efficient resource usage
    - EBS volumes come with fixed IOPS/size ratio
    - Other options (io1) can cost much more under heavy load
      - Cost structure: price per GB-mon + per provisioned IOPS-month
  - Multi Availability Zone (AZ) storage
  - Thin slicing of the storage
    - Some workloads require very small capacity from the storage, using native cloud storage can cost a lot.
The Public Cloud Structure

- Public Clouds
  - AWS, Azure, GCP, IBM, others

- Regions
  - Separate geographic areas for the infrastructure
    - Examples us-east (N. Virginia, Ohio), APAC (Mumbai, Seoul and 4 more), 5 European regions and more

- Availability Zones
  - Multiple data centers within the same region
    - Typically 3, but could be 2 up to 6 (today)
    - For 2 AZ region, there is no real protection from AZ failure, unless we have arbiter outside the region.

Examples use AWS terminology
Examples use AWS terminology - https://storageio.com/images/SIO_AWS_Regions.gif
The Public Cloud Economy

- Replace Capex with Opex
- Pay for use of *everything*
  - **Compute**
    - Node / VM vs. Serverless
  - **Storage**
    - Different storage types with various cost structures
  - **Network**
    - Intra-AZ network is cheaper than Inter-AZ network
  - **Applications**
    - Databases, Dev tool, AI tools, Replications all provided in XaaS model

Efficient public cloud applications minimize resource usage even when the resources exist.
Reducing Storage Cost
TIP 1 - Device Selection

- Use instance storage instead of EBS
  - Cheaper
  - More performant
  - Less reliable
    - So what - storage systems know how to handle unreliable media!

- If there are requirements to use EBS use it sparingly
  - Consider using less replicas or EC schemes
TIP 2 - Network Optimization

- Minimize cross AZ communications
  - Implement AZ read affinity
  - Use compression where appropriate
- Consider alternative topologies
- If appropriate consider using gateways instead of point to point communications
- Save intra-AZ network if the cost for saving is low
Tip 3 - Elasticity / Right Sizing

- Always try to right-size the system
  - Empty system is bad use of money!
- Use relatively small building blocks
  - In order to have tight scaling options
  - Optimize cost/size per your use-case or workload
Tip 4 - Enable Downsizing

- No real support from public cloud providers
  - It is impossible to reduce the size of EBS volumes
  - Therefore, the units of sizing are EBS volumes, or storage nodes - plan the size of these units carefully.
- Enable downsizing for your customers
  - If the customers pay per storage capacity, they will thank you!
Summary - Key Points for Cloud Systems

- **Elasticity**
  - While this is the obvious it is also the key for successful cloud implementation.

- **Multiple deployment options**
  - Pricing model can change without notice, system should be flexible enough to adapt quickly.

- **Think OPEX**
  - Need to optimize on more dimensions
Questions
Thank you.

jsalomon@redhat.com

linkedin.com/company/red-hat
facebook.com/redhatinc
youtube.com/user/RedHatVideos
twitter.com/RedHatLabs